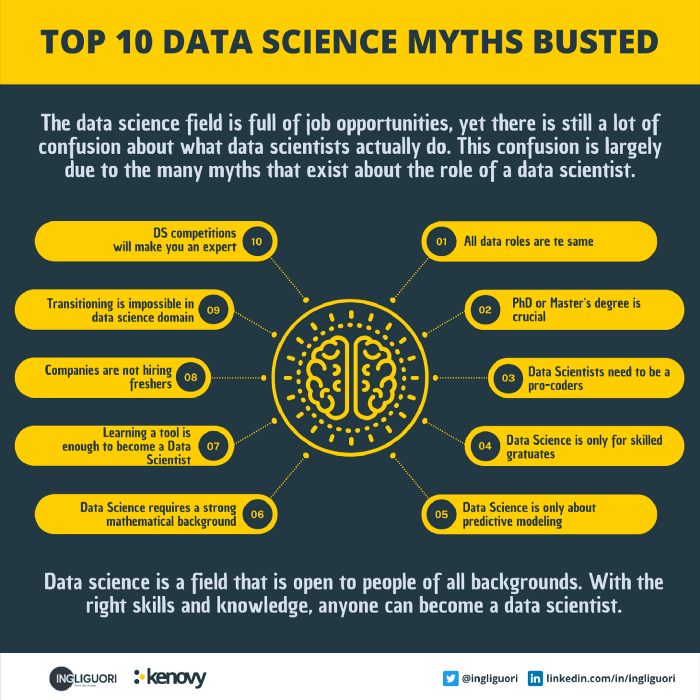
**Top 10 Data Science Myths Busted**

**The data science field is full of job opportunities, yet there is still a lot of confusion about what data scientists actually do. This confusion is largely due to the many myths that exist about the role of a data scientist. In this article, we will bust the top 10 myths about data science. By the end of this article, you will have a better understanding of the role…**

With the popularity of data science, there are a lot of myths about it that circulate around. If you’re interested in pursuing a career in data science, it’s important to be aware of these myths and bust them.

A Big Data World is full of so many job profiles, including data engineers, data scientists, data analysts, business analysts, and many more. There is no doubt that data scientists are the most popular job profile in the field, which is why there is so much confusion around it, especially among beginners. When they try to explore this job profile, they get confused because of the random content available on the Internet that is very confusing, they are unable to decide whether this field is a good fit for them, and they are unable to decide the resources and why this is happening because of random postings and random myths across the globe for the data science profile and that is why we will be busting those top 10 myths related to data science and will help you to debunk them.



Infographic by the author

**Myth 1. All data roles are the same**

Data analysts, data engineers and data scientists are all doing the same thing, which is completely wrong because their jobs, roles and responsibilities are all very different. We understand this creates confusion because all of these people work under the same big data umbrella. First, let’s see what data engineers do. Their responsibility is to work on the basic engineering stuff and build those scalable data pipelines so that raw data can be extracted from multiple sources, transformed, and dumped into downstream systems. Data scientists and data analysts rely on this process because they transform data into something that matters — the process of transforming data into information. That is the key to representing meaningful insight in data and constructing accurate machine-learning models for data scientists. While these people should appear to be doing the same thing, their skills may overlap, but their duties are fundamentally distinct, which is why companies recruit people for different job profiles. If not, they would have hired a data specialist who does it all.

**Myth 2. Ph.D. or Master’s Degree is required to become a Data Scientist**

This is also completely wrong, however, it also depends a lot on the type of job role you want to get. If we are looking for a role in research, for example, we need a master’s or Ph.D., but if you want to solve complex data problems and work around deep learning or machine learning, then you will be the one to work on the data science stuff using the libraries and using the data analysis approaches in order to create those complex data mysteries. Therefore, you don’t need to hold a Master’s Degree. Nowadays everything is all about skills so if you have the right skill set for the data scientist then you can definitely get into it.

**Myth 3. Data Scientists need to be a pro coder**

This is also completely wrong because as a data scientist your job is to work on data extensively and when we talk about pro coding it means working too much on the competitive programming side or having a very in-depth knowledge of typical data structures and algorithms. Surely a data scientist must have good complex problem-solving skills and in the world of data science, we have languages ​​like Python and R which provide very important support through multiple libraries that can be used to solve complex data problems. Your goal as a data scientist should be to understand how to use those libraries and their modules so that you can create the best data models and models related to machine learning. The data scientist must have these skills in order not to waste too much time on competitive programming or practicing typical data structures and algorithms.

**Myth 4. Data Science is only for CS graduates**

This is one of the most important myths that must be debunked. Although more and more young people decide to study science subjects, mainly attracted by the growing offer of jobs in the technological field, to date the majority of people do not have a technical background. Companies tend to hire candidates for the data science profile, and those with non-tech and non-technical backgrounds are selected because they are really good at problem-solving and understanding business use cases. These are the important things to understand in order to be successful in data science interviews. Companies do not focus on the typical technical skills of a programmer, but they want to understand if the candidates are good in the aptitude part, if they are able to understand the case studies, if they are able to understand what kind of business matrices they can extract from the data and what kinds of complex data-related problems they can solve. So never think that coming from a non-CS or non-tech background you cannot enter the data science profile. By the way for CS graduates, their first priority is still software engineering and they want to make that career out of their career. That is why in this domain of data science, there are more and more job opportunities for non-tech people. So nowadays it doesn’t matter if you have a specific degree related to CS, you will still have the chance to work as a data scientist if you can acquire the right skills for the data science profile.

**Myth 5. Data Science is only about predictive modeling**

Not everyone knows that data scientists invest 80% of their time in cleaning and transforming data and 20% of their time is dedicated to the data modeling part, so a data scientist who wants to create very accurate data as well as a machine learning model, he needs to clean and transformed data. We know that when we work on a particular Big Data solution, there are multiple steps involved in it and the first and very important part is transforming the data. Nowadays, we receive data from multiple sources, and the raw data sometimes contains errors, as well as junk records. If we cannot clean our data, we will not be able to obtain meaningful transformation data, and we will not be able to create machine learning models that are very accurate. That’s why data science is not just about building the predictive model and regression models, it’s a good mix of cleaning and transforming the data, and then building accurate machine learning models.

**Myth 6. Data Science requires a strong mathematical background**

This is also completely not true because being good at math is one of the important parts of your day-to-day activities as a data scientist. While analyzing the data we would need those mathematical concepts, let’s say aggregating the data statistics part, probability part, but that is something which is and not a must to have a skill set to become a data scientist. As we know in data science, we have amazing programming languages like Python and R which provides great support for amazing libraries as well which we can directly use in order to perform that typical mathematical calculation and computation so unless you need to innovate something, or you need to create any new algorithm you don’t need to be a mathematics expert.

**Myth 7. Learning a tool is enough to become a Data Scientist**

A good data science profile is a combination of multiple skills, technical skills plus non-technical skills. Both are required to become good data scientists. In order to become a good data scientist, you cannot just rely on programming or any specific tool you think is being used in data science. If you are a good expert in both stuff, you are likely to become a good data scientist. As data scientists, we have to interact with multiple stakeholders when working on complex data problems, and we must work directly with the business to gather all the requirements. Understanding the data domain, why we are processing it, what insights we can gain from the transform data, how to solve it, and what is relevant and what is irrelevant, are all necessary for this data science domain. Don’t think that only technical stuff or any specific tool like a language or a database is just enough to crack the data science profile. You need to use your non-technical skills and soft skills as your support system to become a good data scientist.

**Myth 8. Companies are not hiring freshers for the data science role**

So this is a common question. If we would have discussed this topic five-six years back, then this was completely true. Companies were not focusing on the freshers for the data science role, but now in 2022, this has changed a lot because freshers nowadays are self-aware and self-motivated, and they are very interested in exploring fields such as data science, and data engineering. Instead of depending on others, they are putting their own effort into exploring these positions. They are also taking participation actively in competitions in hackathons open source contributions and trying to build cool projects on their own. That’s how they are able to acquire the right skill sets for the data science profile and also amazing dev skills. That’s why now companies are preferring freshers to join them to cope with the vacancies in data science roles.

**Myth 9. Transitioning into data science domain is impossible from the different job profile**

If you come from a data-related background, let’s say you worked as a data engineer, data analyst, or business analyst then this transition will be easy for you because you already know about the data and how to work with it different tools and tech frameworks. On the other hand, if you are coming from a different job profile as a testing profile or a software engineering profile then transitioning into data science will be challenging but not impossible. You need to work on a skill set you need to acquire and then work on some realistic projects that can help you in acquire new capabilities and develop your interest in the data science domain to understand how you really can add value to the company as a data scientist.

**Myth 10. Data Science competitions are going to make you an Expert**

Data science competitions are a good idea to acquire the right skills and also to get an idea of the data science environment and even acquire development skills, but if you think that by only participating in hackathons, and creating projects in competitions, you can help you become a data science expert then this is not true. You will improve and add value to your resume, so that you can showcase your work and what kind of results you have achieved, but if you really want to be an expert you need to work on some realistic use cases or a production-level application. For this reason, it is advisable that freshmen are directed toward a good internship.

**Conclusions**

If you’re interested in data science, don’t let these myths discourage you. Data science is a field that is open to people of all backgrounds. With the right skills and knowledge, anyone can become a data scientist.